

Deriving the Hirshfeld Partitioning Using Distance Metrics

Farnaz Heidar-Zadeh¹, Paul W. Ayers^{1*}, Patrick Bultinck²

1. Dept. of Chemistry & Chemical Biology; McMaster University; Hamilton, Ontario, Canada

2. Dept. of Physical and Inorganic Chemistry; Ghent University; Ghent, Belgium

The atoms in molecules associated with the Hirshfeld partitioning minimize the generalized Hellinger-Bhattacharya distance to the reference pro-atom densities. Moreover, the reference pro-atoms can be chosen by minimizing the distance between the pro-molecule density and the true molecular density. This provides an alternative to both the heuristic “stockholder” and the mathematical information-theoretic interpretations of the Hirshfeld partitioning. These results extend to any member of the family of f -divergences.

Nalewajski and Parr provided the impetus for recent work on information-theoretic approaches to population analysis.^{1-33,34} Specifically, they derived the atom-in-a-molecule densities, $\rho_A(\mathbf{r})$, that are as close as possible to the reference pro-atom densities, $\rho_A^0(\mathbf{r})$, in an information-theoretic sense, subject to the constraint that the sum of the atomic densities exactly recovers the molecular density, $\rho_{\text{mol}}(\mathbf{r})$.¹ I.e.,

$$\rho_A(\mathbf{r}) = \arg \min_{\left\{ \rho_A(\mathbf{r}) \left| \rho_{\text{mol}}(\mathbf{r}) = \sum_{A=1}^{N_{\text{atoms}}} \rho_A(\mathbf{r}) \right. \right\}} \sum_{A=1}^{N_{\text{atoms}}} \int \rho_A(\mathbf{r}) \ln \left(\frac{\rho_A(\mathbf{r})}{\rho_A^0(\mathbf{r})} \right) d\mathbf{r} \quad (1)$$

The minimization can be performed analytically, giving the Hirshfeld partitioning,³⁵

* ayers@mcmaster.ca

$$\rho_A(\mathbf{r}) = \left(\frac{\rho_A^0(\mathbf{r})}{\sum_{B=1}^{N_{\text{atoms}}} \rho_B^0(\mathbf{r})} \right) \rho_{\text{mol}}(\mathbf{r}) = w_A(\mathbf{r}) \rho_{\text{mol}}(\mathbf{r}). \quad (2)$$

While Hirshfeld chose isolated neutral pro-atoms, subsequent work has considered more general pro-atoms.²⁵⁻³¹ However, because the information measure in Eq. (1) is generally applied only to normalized probability distribution functions, many researchers constrain the pro-atoms to have the same charge as the corresponding atoms-in-a-molecule,

$$\int \rho_A(\mathbf{r}) d\mathbf{r} = \int \rho_A^0(\mathbf{r}) d\mathbf{r}. \quad (3)$$

The iterative Hirshfeld method (Hirshfeld-I)²⁵ is one of many examples of this type of partitioning. While the constraint in Eq. (3) is not strictly necessary, methods that avoid this constraint are complicated by additional “entropy of mixing” terms.^{3,24}

The information-theoretic derivation is disadvantaged by the fact that the Kullback-Leibler divergence

$$I_{\text{KL}}(\rho|\rho^0) = \int \rho(\mathbf{r}) \ln \left(\frac{\rho(\mathbf{r})}{\rho^0(\mathbf{r})} \right) d\mathbf{r} \quad (4)$$

is not a distance metric. However, as we now show, the Hirshfeld atom is also obtained using a generalization of the Hellinger-Bhattacharya distance metric,³⁶

$$d(\rho, \rho^0) = \int \left((\rho(\mathbf{r}))^{1/n} - (\rho^0(\mathbf{r}))^{1/n} \right)^2 d\mathbf{r} \quad n \neq \{0, 1\} \quad (5)$$

in place of the Kullback-Leibler divergence in Eq. (1). Here we allow n to be any real number except zero and one. The conventional Hellinger-Bhattacharya distance (the square root of the $n = 2$ case) is commonly used in density functional theory,^{37,38} where it is among the simplest reasonable measures for the distance between electron densities.

The atom-in-molecule densities are obtained by differentiating the Lagrangian,

$$\Lambda[\{\rho_A\}] = \sum_{A=1}^{N_{\text{atoms}}} \int \left((\rho_A(\mathbf{r}))^{1/n} - (\rho_A^0(\mathbf{r}))^{1/n} \right)^n d\mathbf{r} - \int \lambda(\mathbf{r}) \left(\rho_{\text{mol}}(\mathbf{r}) - \sum_{A=1}^{N_{\text{atoms}}} \rho_A(\mathbf{r}) \right) d\mathbf{r} \quad (6)$$

obtaining the equations,

$$0 = \frac{\delta \Lambda[\{\rho_A\}]}{\delta \rho_A(\mathbf{r})} = \frac{(\rho_A(\mathbf{r}))^{1/n} - (\rho_A^0(\mathbf{r}))^{1/n}}{(\rho_A(\mathbf{r}))^{1/n}} + \lambda(\mathbf{r}). \quad (7)$$

Therefore the ratio of the pro-atomic density to the atom-in-molecule density is the same for all atoms,

$$\frac{\rho_A^0(\mathbf{r})}{\rho_A(\mathbf{r})} = (1 + \lambda(\mathbf{r}))^n \quad (8)$$

Using the constraint that the atomic densities sum to the total density results in the Hirshfeld partitioning, (2).^{1,23} This analysis can be extended to include the $n \rightarrow 0$ limit. In the $n = 0$ case, there are many possible partitionings, but the Hirshfeld partitioning is a particular solution.

The Hirshfeld partitioning, Eq. (2), is thus derived for any choice of pro-atom densities. The “best” choice of pro-atom densities could be selected by minimizing the generalized Hellinger-Bhattacharya distance with respect to the pro-atoms,

$$\min_{\{N_A^0\}} \min_{\left\{ \rho_A(\mathbf{r}) \left| \rho_{\text{mol}}(\mathbf{r}) = \sum_{A=1}^{N_{\text{atoms}}} \rho_A(\mathbf{r}) \right. \right\}} \sum_{A=1}^{N_{\text{atoms}}} \int \left((\rho_A(\mathbf{r}))^{1/n} - (\rho_A^0(\mathbf{r}, N_A^0))^{1/n} \right)^n d\mathbf{r}, \quad (9)$$

where, for simplicity, we have assumed that only the number of electrons in the isolated pro-atom densities, N_A^0 , is a variational parameter. (If desired, electron densities of atoms with fractional electron number should be constructed by linear interpolation between atomic

densities with integer electron number.³⁹⁻⁴¹) Now, using the explicit formula for the atom-in-a-molecule density,

$$\begin{aligned}
& \min_{\{N_A^0\}} \sum_{A=1}^{N_{\text{atoms}}} \int \left(\left(\frac{\rho_A^0(\mathbf{r}, N_A^0) \rho_{\text{mol}}(\mathbf{r})}{\sum_{B=1}^{N_{\text{atoms}}} \rho_B^0(\mathbf{r}, N_B^0)} \right)^{1/n} - \left(\rho_A^0(\mathbf{r}, N_A^0) \right)^{1/n} \right)^n d\mathbf{r} \\
&= \min_{\{N_A^0\}} \int \sum_{A=1}^{N_{\text{atoms}}} \left(\frac{\rho_A^0(\mathbf{r}, N_A^0)}{\sum_{B=1}^{N_{\text{atoms}}} \rho_B^0(\mathbf{r}, N_B^0)} \right) \left(\left(\rho_{\text{mol}}(\mathbf{r}) \right)^{1/n} - \left(\sum_{B=1}^{N_{\text{atoms}}} \rho_B^0(\mathbf{r}, N_B^0) \right)^{1/n} \right)^n d\mathbf{r} \quad (10) \\
&= \min_{\{N_A^0\}} \int \left(\left(\rho_{\text{mol}}(\mathbf{r}) \right)^{1/n} - \left(\sum_{B=1}^{N_{\text{atoms}}} \rho_B^0(\mathbf{r}, N_B^0) \right)^{1/n} \right)^n d\mathbf{r}
\end{aligned}$$

That is, the optimal choice for the pro-atom densities is the one that minimizes the discrepancy between the molecular density and the pro-molecular density,

$$\rho_{\text{mol}}^0(\mathbf{r}) = \sum_{A=1}^{N_{\text{atoms}}} \rho_A^0(\mathbf{r}, N_A^0). \quad (11)$$

One could add additional constraints on the optimization of the pro-atoms. For example, it is probably advisable to force the pro-molecule and molecule have the same charge. Alternatively, one could constrain the pro-atoms and the atoms to have the same charge, although this is no longer essential because we are not using information theory. This is fortunate, since (I) minimizing the Kullback-Leibler divergence with the $N_A^0 = N_A$ constraint gave disappointing results²⁶ and (II) constraints that involve both the pro-atomic and atomic densities couples the inner and outer minimization in Eq. (9), which complicates the numerical analysis, as Eq. (2) is no longer valid. Therefore, if one wishes to impose the $N_A^0 = N_A$

constraint, it seems preferable to choose the optimal pro-atom nonvariationally, as is done, for example, in Hirshfeld-I.²⁵

Both the Hirshfeld partitioning itself, Eq. (2), and the equivalence between optimizing the pro-atoms and the pro-molecule, Eq. (10), turn out to be very general. It is straightforward to demonstrate that these results hold for any member of the family of Csiszár-Morimoto-Ali-Silvey f -divergences,⁴²⁻⁴⁴

$$\sum_{A=1}^{N_{\text{atoms}}} \int \rho_A(\mathbf{r}) f\left(\frac{\rho_A^0(\mathbf{r})}{\rho_A(\mathbf{r})}\right) d\mathbf{r} \quad (12)$$

where $f(x)$ is *any* function that is convex for $x > 0$ and has $f(1) = 0$.⁴⁵ This family encompasses not only the Hellinger-Bhattacharya distance (using $f(x) = (\sqrt{x} - 1)^2$) and the Kullback-Leibler information (using $f(x) = -\ln x$), but also many of the other information measures that have been used to derive the Hirshfeld partitioning. For example, the Tsallis information measure in ref. ²⁴ is proportional to the α -divergence, $f(x) = 4(1 - x^{(1+\alpha)/2}) / (1 - \alpha^2)$. Symmetrized forms, where the pro-atoms and atoms enter the formula in identical fashion,²³ can be obtained by replacing any function $f(x)$ by $g(x) = f(x) + x f(1/x)$.

In the original work of Nalewajski and Parr,¹ the Hirshfeld partitioning was shown to minimize the Kullback-Leibler divergence between an atom and its reference pro-atom. Here we show that the Hirshfeld partitioning minimizes the Hellinger distance between the atom-in-a-molecule density and the pro-atom density. (Even more generally, the Hirshfeld partitioning minimizes the f -divergence between the atom-in-a-molecule density and the pro-atom density.) Deriving the Hirshfeld partitioning from a distance measure makes it clear that it is not essential to choose the reference pro-atom densities and the atom-in-a-molecule densities to have the same

population. This suggests that this commonly used constraint²⁵⁻²⁷ can be relaxed, and that the pro-atom densities can be selected by minimizing Hellinger-Bhattacharya distance between the atomic and the pro-atom densities. This is equivalent to choosing the pro-atoms so that the pro-molecule density is as close as possible to the molecular density.

Acknowledgement: FH-Z was supported by a Vanier-CGS fellowship and a Michael Smith Foreign Study Supplement from the National Sciences and Engineering Research Council of Canada (NSERC); PWA was supported by a Discovery Grant and a E. W. R. Steacie Fellowship from NSERC; PB was supported by the Flemish Research Foundation (FWO-Vlaanderen). The authors thank Toon Verstraelen for many helpful discussions on information-theoretic atoms in molecules.

References

- [1] R. F. Nalewajski, and R. G. Parr, Proc. Natl. Acad. Sci. **97**, 8879 (2000).
- [2] R. F. Nalewajski, and R. G. Parr, J. Phys. Chem. A **105**, 7391 (2001).
- [3] R. G. Parr, P. W. Ayers, and R. F. Nalewajski, J. Phys. Chem. A **109**, 3957 (2005).
- [4] R. F. Nalewajski, Adv. Quantum Chem. **43**, 119 (2003).
- [5] R. F. Nalewajski, J. Phys. Chem. A **104**, 11940 (2000).
- [6] R. F. Nalewajski, and E. Broniatowska, J. Phys. Chem. A **107**, 6270 (2003).
- [7] R. F. Nalewajski, Adv. Quantum Chem. **51**, 235 (2006).
- [8] R. F. Nalewajski, Int. J. Quantum Chem. **108**, 2230 (2008).
- [9] S. B. Liu, J. Chem. Phys. **126**, 191107 (2007).
- [10] A. Nagy, and S. B. Liu, Phys. Lett. A **372**, 1654 (2008).
- [11] R. O. Esquivel, S. B. Liu, J. C. Angulo, J. S. Dehesa, J. Antolin, and M. Molina-Espiritu, J. Phys. Chem. A **115**, 4406 (2011).
- [12] S. B. Liu, J. Chem. Phys. **126**, 244103 (2007).
- [13] A. Nagy, Chem. Phys. Lett. **449**, 212 (2007).
- [14] F. De Proft, R. Vivas-Reyes, A. Peeters, C. Van Alsenoy, and P. Geerlings, J. Comput. Chem. **24**, 463 (2003).
- [15] R. F. Nalewajski, and E. Switka, PCCP **4**, 4952 (2002).
- [16] R. F. Nalewajski, E. Switka, and A. Michalak, Int. J. Quantum Chem. **87**, 198 (2002).
- [17] R. F. Nalewajski, J. Phys. Chem. A **107**, 3792 (2003).
- [18] P. Bultinck, P. W. Ayers, S. Fias, K. Tiels, and C. Van Alsenoy, Chem. Phys. Lett. **444**, 205 (2007).
- [19] S. Van Damme, P. Bultinck, and S. Fias, J. Chem. Theory Comp. **5**, 334 (2009).

- [20] D. E. P. Vanpoucke, P. Bultinck, and I. Van Driessche, *J. Comput. Chem.* **34**, 405 (2013).
- [21] D. Vanfleteren, D. Van Neck, P. Bultinck, P. W. Ayers, and M. Waroquier, *J. Chem. Phys.* **133**, 231103 (2010).
- [22] D. Vanfleteren, D. Van Neck, P. Bultinck, P. W. Ayers, and M. Waroquier, *J. Chem. Phys.* **136**, 014107 (2012).
- [23] P. W. Ayers, *J. Chem. Phys.* **113**, 10886 (2000).
- [24] P. W. Ayers, *Theor. Chem. Acc.* **115**, 370 (2006).
- [25] P. Bultinck, C. Van Alsenoy, P. W. Ayers, and R. Carbó-Dorca, *J. Chem. Phys.* **126**, 144111 (2007).
- [26] D. Ghillemijn, P. Bultinck, D. Van Neck, and P. W. Ayers, *J. Comput. Chem.* **32**, 1561 (2011).
- [27] T. Verstraelen, P. W. Ayers, V. Van Speybroeck, and M. Waroquier, *J. Chem. Theory Comp.* **9**, 2221 (2013).
- [28] T. C. Lillestolen, and R. J. Wheatley, *Chem. Commun.* **45**, 5909 (2008).
- [29] T. Verstraelen, P. W. Ayers, V. Van Speybroeck, and M. Waroquier, *Chem. Phys. Lett.* **545**, 138 (2012).
- [30] T. A. Manz, and D. S. Sholl, *J. Chem. Theory Comp.* **8**, 2844 (2012).
- [31] T. A. Manz, and D. S. Sholl, *J. Chem. Theory Comp.* **6**, 2455 (2010).
- [32] F. Heidar-Zadeh, P. Fuentealba, C. Cardenas, and P. W. Ayers, *PCCP* **16**, 6019 (2014).
- [33] S. Liu, C. Rong, and T. Lu, *J. Phys. Chem. A* **118**, 3698 (2014).
- [34] P. Bultinck and D. Jayatilaka, *Comp. Theor. Chem.* (accepted). DOI: 10.1016/j.comptc.2014.06.017
- [35] F. L. Hirshfeld, *Theor. Chim. Act.* **44**, 129 (1977).

- [36] M. M. Deza, and E. Deza, *Encyclopedia of Distances* (Springer-Verlag, Berlin, 2013), 2 edn.
- [37] E. H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).
- [38] H. Eschrig, *The Fundamentals of Density Functional Theory* (Eagle, Leipzig, 2003).
- [39] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982).
- [40] P. W. Ayers, J. Math. Chem. **43**, 285 (2008).
- [41] W. T. Yang, Y. K. Zhang, and P. W. Ayers, Phys. Rev. Lett. **84**, 5172 (2000).
- [42] S. M. Ali, and S. D. Silvey, Journal of the Royal Statistical Society Series B-Methodological **28**, 131 (1966).
- [43] T. Morimoto, J. Phys. Soc. Jpn. **18**, 328 (1963).
- [44] I. Csiszár, Magyar. Tud. Akad. Mat. Kutato Int. Kozl **8**, 85 (1963).
- [45] F. Heidar-Zadeh, P. W. Ayers, in preparation.